Supporting Information

Mets and Brainard 10.1073/pnas.1713031115

SI Materials and Methods

Audio Recording and Initial Processing. For audio recording birds were single-housed in sound-isolation chambers (Acoustic Systems). Songs were recorded digitally using custom Python or LabView (National Instruments) software at a sampling frequency of 32 kHz, a bit depth of 16, and stored uncompressed. Recording microphones were placed in a fixed position at the top of the cage housing the bird. Before further analysis, all songs were digitally high-pass-filtered at ~500 Hz using a digitally implemented elliptical infinite impulse response filter with a passband edge frequency of 0.04 rad. All recordings used for analysis were acquired during early adulthood (90–120 d posthatch).

Syllable Segmentation and Song Tempo Calculation. Song tempo was quantified as the average number of syllables produced per second of song. Discrete units of sound separated by silence (syllables) were identified based on amplitude. First an "amplitude envelope" was created by rectifying the song waveform and then smoothing the waveform through convolution with an 8-ms square wave. Threshold crossings of this amplitude trace were then used to identify periods of vocalization. Thresholds were set heuristically to result in segmentation that corresponded with syllable onsets and offsets apparent in examination of spectrograms. Once the threshold was established, "objects" were identified as uninterrupted regions of the amplitude envelope over threshold longer than 10 ms. Any objects separated by a gap of 5 ms or less were merged, producing a final set of objects that were defined as syllables. A series of syllables that had no gaps larger than 250 ms was considered a song bout. For each bird, tempo was then quantified as the number of syllables present in a song bout divided by the duration of the song bout, averaged across at least 60 bouts.

Heritability Calculations. The genetic contribution of the father was calculated as the father–offspring regression slope as estimated by the method of OLS (1) both when estimated alone and when estimated jointly with experiential influences. When an interaction term was included, the weights presented are "centered" in that they are the estimated weight of each parameter given average values for all parameters.

Statistics. Sample sizes were selected to detect heritability estimates of a minimum of ~10%. No animals were removed from the study. As no subjective measurements were made, no blinding was performed. When used, statistical tests were appropriate to the data presented and the data, to the limit of detection, were consistent with the assumptions of the tests. When appropriate, variability is reported.

1. Lynch M, Walsh B (1998) *Genetics and Analysis of Quantitative Traits* (Sinauer, Sunderland, MA).

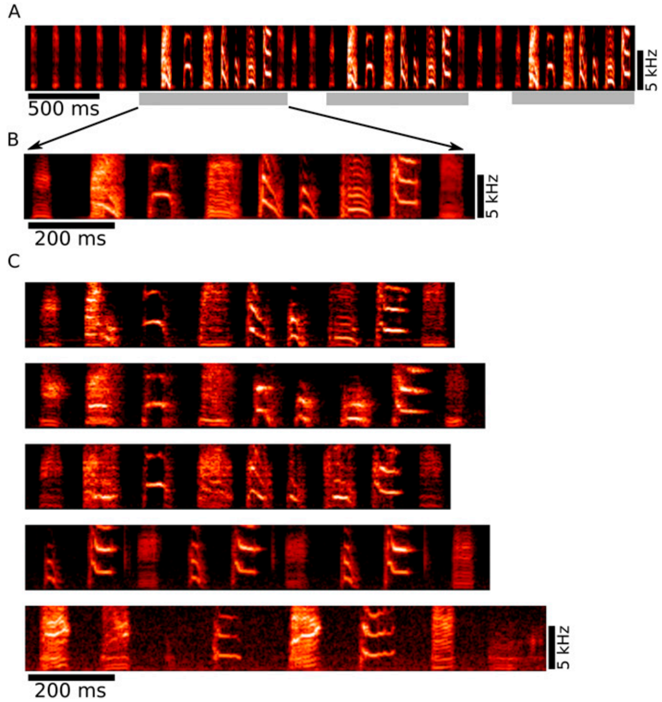

Fig. S1. Computer tutor-driven song learning. (*A*) Spectrogram illustrating the 8.5-syl/s synthetic tutor song stimulus used for computer tutoring. The tutor song was composed of nine categorically distinct syllables from our database of recorded Bengalese finch songs. Syllables were chosen to represent a range of syllable types found in normal Bengalese finch song. The intersyllable gap durations were drawn from the distribution of intersyllable gaps present in our colony. Each identical playback event was composed of a series of introductory syllables followed by three repetitions of a stereotyped sequence of syllables, or motif (gray bars). (*B*) An enlarged spectrogram of the tutor song motif. (*C*) Spectrograms of song segments that qualitatively most closely resemble the tutor song motif from five computer-tutored birds. These illustrate a typical range in the quality of tutor song copying and are arranged from very good (top) to comparatively poor (bottom). Song tempos were measured as syllables per second averaged across multiple song bouts (*Methods*).